

Estimating the Empirical Probability of Submarine Landslide Occurrence

E.L. Geist and T. Parsons

Abstract The empirical probability for the occurrence of submarine landslides at a given location can be estimated from age dates of past landslides. In this study, tools developed to estimate earthquake probability from paleoseismic horizons are adapted to estimate submarine landslide probability. In both types of estimates, one has to account for the uncertainty associated with age-dating individual events as well as the open time intervals before and after the observed sequence of landslides. For observed sequences of submarine landslides, we typically only have the age date of the youngest event and possibly of a seismic horizon that lies below the oldest event in a landslide sequence. We use an empirical Bayes analysis based on the Poisson-Gamma conjugate prior model specifically applied to the landslide probability problem. This model assumes that landslide events as imaged in geophysical data are independent and occur in time according to a Poisson distribution characterized by a rate parameter λ . With this method, we are able to estimate the most likely value of λ and, importantly, the range of uncertainty in this estimate. Examples considered include landslide sequences observed in the Santa Barbara Channel, California, and in Port Valdez, Alaska. We confirm that given the uncertainties of age dating that landslide complexes can be treated as single events by performing statistical test of age dates representing the main failure episode of the Holocene Storegga landslide complex.

Keywords Submarine landslides • age dating • probability

E.L. Geist (✉) and T. Parsons

U.S. Geological Survey, 345 Middlefield Road, MS 999, Menlo Park, CA 94025, USA

e-mail: egeist@usgs.gov

1 Introduction

Estimating the likelihood of submarine landslide occurrence is an important component of an overall hazard assessment. This likelihood applies not only to the direct hazard posed by this geologic phenomenon to offshore structures, pipelines, and submarine cables but also to the indirect hazard of tsunami generation. For the latter, knowledge of both the distribution of landslide sizes and the distribution of events in time is necessary to comprehensively quantify tsunami hazards at a coastline (Geist et al. 2009; Geist and Parsons 2009). The distribution of submarine landslide sizes has been suggested to follow a power-law, in the case of debris avalanches off the north coast of Puerto Rico (ten Brink et al. 2006) or a log-normal distribution in the case of debris flows off the U.S. Atlantic coast (Chaytor et al., 2009). In this paper, we instead focus on methods to determine the rate of occurrence of submarine landslides in time, where recurrent landslides are embedded in a well-developed stratigraphic sequence. We also limit our analysis to empirical techniques. Computational techniques have been developed to determine landslide probability based on the probability distribution of the triggering mechanism (i.e., strong seismic ground motion for most of the landslides considered) (ten Brink et al., 2009) and based on long-term stratigraphic simulations (Hutton and Syvitski 2004). However, there is significant uncertainty in determining the probability of slope failure conditioned upon not only the triggering mechanism but also geotechnical parameters such as the factor of safety and distribution of pore pressure. Where data exist on the past occurrence of submarine landslides, we argue that uncertainty in occurrence rates is better quantified by empirical techniques.

The objective of determining probability is obviously hampered by the lack of an instrumental record of submarine landslide occurrence (as we have for submarine earthquakes) and the sparse number of age dates from core data relative to the number of submarine mass movements that have been mapped throughout the world. However, we can avail ourselves of probabilistic techniques that have been previously developed to estimate earthquake occurrence rates from onshore paleoseismic sequences (i.e., from excavation trenches across faults or from coastal records of coseismic subsidence/uplift). The problem of predicting the occurrence rate of both earthquakes and landslides from paleo-events share some similarities: in determining the occurrence rates, one has to take into account the open time intervals (before the first and after the last event in the sequence) as well as the uncertainty in age dating the horizons. For the paleoseismic problem, both Bayesian and Monte Carlo techniques have been developed (Savage 1994; Ogata 1999; Parsons 2008). One difference between observed submarine landslide sequences and onshore paleoseismic sequences is that for landslides, we often only have age dates of the most recent event and the age date from a horizon that underlies the landslide sequence. After a brief description of the field data used to identify landslide sequences, we introduce an empirical Bayes analysis particularly suited for this case. The empirical Bayes analysis is based on a Poisson distribution that assumes landslide events are independent from one to another. Finally, we examine whether landslide complexes

such as Storegga, in the Norwegian Sea, can be treated as a single event for the empirical Bayes analysis in Section 4.

2 Data

Two types of data are needed for this analysis: (1) geophysical data to both map landslides at the sea floor and to identify previous occurrences of landslides in cross section and (2) age dates of sediment that bracket the landslide event. Multibeam bathymetry and side scan sonar technologies have successfully been used to map landslide complexes in great detail. In cross-section, older landslides are identified from seismic reflection data using a variety of controlled sources, depending on the desired resolution and depth of acoustic penetration.

To date a landslide, ideally age dates from pre- and post-landslide rocks are obtained from cores along its distal edge (e.g., Normark et al. 2004). It is also possible to tie seismic reflectors above and below the landslide to ages from nearby drill holes (e.g., Fisher et al. 2005). The method of dating is typically radiocarbon analysis of microfossils, though other techniques are also used (e.g., Lee 2009). Dating submarine landslides is complicated by uncertainty that arises from the large carbon reservoir in the oceans, which causes radiocarbon ages of marine samples to appear several hundred years older than coeval terrestrial samples to globally varying degrees (Stuiver and Braziunas 1992). It is rare to have dates for all landslides in an observed sequence – typically we may only have the date of the youngest landslide and the date of a seismic horizon that underlies the landslide sequence.

3 Empirical Bayes Analysis

Bayesian methods are particularly well suited to estimate the empirical probability and associated uncertainty from small sample sizes. To start with, we assume that landslide events occur in time according to a Poisson distribution, such that the number of events n occurring in time t is distributed according to

$$p(n|\lambda, t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!} \quad (1)$$

where λ is the rate or intensity parameter of the distribution. This distribution is commonly used for phenomena that are rare, but have very many opportunities to happen. One property of the Poisson distribution is that the time between events (inter-event times) are independently distributed. In addition, in its cumulative form, the probability that one or more events will occur in a particular time window does not vary with the time since the last event.

For many observed submarine landslide sequences, the dates of individual landslide events are unknown. However landslides occur in a stratigraphic sequence, and commonly, a datable horizon exists below the sequence. The rate parameter λ itself is uncertain and can be treated as a random variable. To estimate the distribution of values that λ can attain for this type of problem, we can use Bayes theorem, which states that for any parameter θ and set of empirical data $z = z_1, \dots, z_m$, the posterior distribution $\pi(\theta|z)$ is given by

$$\pi(\theta|Z) = \frac{L(Z|\theta)\pi(\theta)}{\int L(Z|\theta)\pi(\theta)d\theta} \quad (2)$$

where $L(z|\theta)$ is the likelihood function, $\pi(\theta)$ is the prior distribution and the integral in the denominator is the marginal probability.

The empirical Bayes approach is greatly simplified if one chooses a prior probability distribution that is “conjugate” to the likelihood function: that is, if the resulting posterior distribution is in the same distribution family (e.g., the exponential distribution family) as the prior distribution. The distribution pairs are often called “conjugate priors”. The conjugate prior to the Poisson distribution is the Gamma distribution, with a probability density function given by

$$f(x) = \frac{\beta^\gamma x^{\gamma-1} e^{-\beta x}}{\Gamma(\gamma)} \quad (3)$$

where the mean is γ/β and the variance is γ/β^2 . Taking the rate parameter λ as the random variable in the prior distribution Eq. 3, it can be shown that with a Poisson likelihood function, the corresponding posterior distribution is also a Gamma distribution (Mortgat and Shah 1979; Campbell 1982) such that

$$\pi(\lambda|z) = \frac{(\beta')^\gamma \lambda^{\gamma'-1} e^{-\beta'\lambda}}{\Gamma(\gamma')} \quad (4)$$

This is often called the Poisson-Gamma model. The posterior distribution parameters γ' and β' are calculated from the number of landslide occurrences (N) and the time period of observation (T):

$$\gamma' = N + (\mu/\sigma)^2 \quad (5)$$

$$\beta' = T + \mu/\sigma^2 \quad (6)$$

where μ and σ are the best estimate of the mean and standard deviation for the rate parameter λ (Mortgat and Shah 1979; Campbell 1982). The resulting posterior probability of the rate parameter is almost identical to that determined from generalized Monte Carlo analysis, such as used in Parsons and Geist (2009), where discrete event intervals were not used. Applying this Poisson-Gamma model to submarine landslide occurrence, we can estimate the most likely mean return time ($1/\lambda$) and

uncertainty through equation Eq. 4. A similar formulation has been developed by Savage (1994) for analyzing paleoseismic horizons, but using the Beta-Binomial conjugate prior model.

3.1 Case Study 1: Santa Barbara Channel

The Bayes Poisson-Gamma model is first applied to a landslide sequence identified in the Santa Barbara Channel, southern California (Fisher et al. 2005) that includes the Goleta landslide complex as its youngest event. Submarine landslides there may present a significant tsunami hazard (Borrero et al. 2001). A seismic reflection profile across the Goleta landslide complex shows several landslide events (Fig. 1). The seismic horizon labeled “A” (yellow line) is estimated to be 170ka old from correlating seismic horizons with an oxygen-isotope curve (red line) at a nearby Ocean Drilling Project hole (Site 893).

For this example, therefore, we take the number of landslide occurrences (N) as three and the time base as 170ka. The resulting distribution of mean return times from the Poisson-Gamma model is shown by the blue curve in Fig. 2.

Fisher et al. (2005) indicate that a seismic horizon B (not shown on Fig. 1) underlies the deepest landslide and is correlated with a 200ka biostratigraphic age

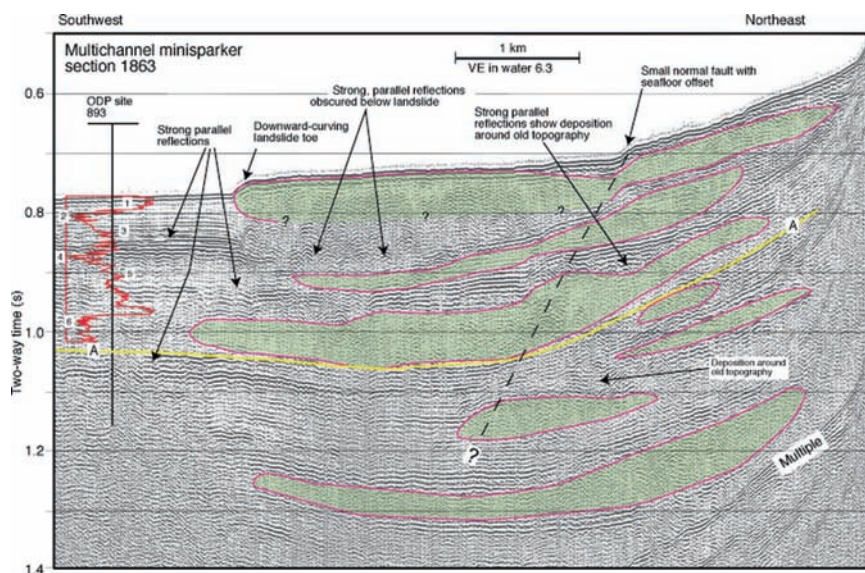


Fig. 1 High-resolution minisparker seismic reflection data oriented parallel to the axis of the Goleta landslide complex. Red outline and green shading indicates interpreted landslide deposits. Red curve is the oxygen isotope variation from ODP site 893. Numbers represent marine isotope stages. Figure from Fisher et al. (2005)

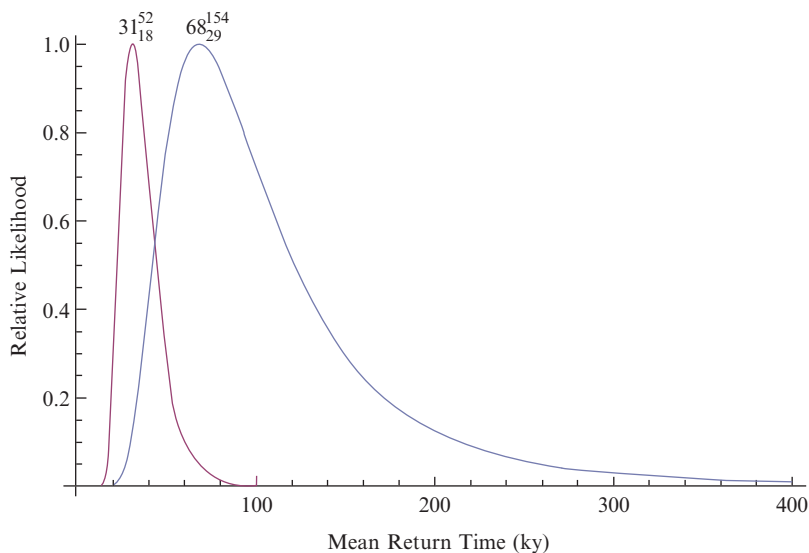


Fig. 2 Distribution of mean return times ($1/\lambda$) for two estimates of the occurrence of landslides in Santa Barbara Channel. Blue curve: three landslide occurrences younger than Horizon A (170 ka) shown in Fig. 1. Red curve: seven landslide occurrences younger than approximately 200 ka. Most likely mean return time (ka) and 95% confidence range (indicated by sub- and superscripts) in the mean return time indicated for each curve

at the nearby ODP hole. The corresponding rate estimate based on all of the imaged landslides and the age of horizon B is shown by the red curve (Fig. 2). It should be noted, however, that this curve includes much smaller landslides than the Goleta landslide complex. In addition, although the red curve appears to have less uncertainty than the blue curve, the 200 ka age used in calculating the red curve may be more uncertain than the age of Horizon A used in the blue curve.

3.2 Case Study 2: Port Valdez

The next case study we examine is landslide occurrence in Port Valdez, Alaska, an enclosed fjord. Ryan et al. (this volume) used seismic reflection data to find that as many as five debris flow events underlie the most recent one, which was triggered by the $M = 9.2$ 1964 Great Alaska earthquake. Hence, in the statistical analysis $N = 6$. The 1964 earthquake in Port Valdez triggered several nearly simultaneous debris flows, one of which destroyed part of the town of Valdez. Other flows generated tsunamis resulting in high tsunami runups around the coastline. The ages of older debris flows are estimated by determining the sedimentation rate in Port Valdez. The first debris flow event before the 1964 debris flow is interpreted to be contemporaneous with the penultimate megathrust

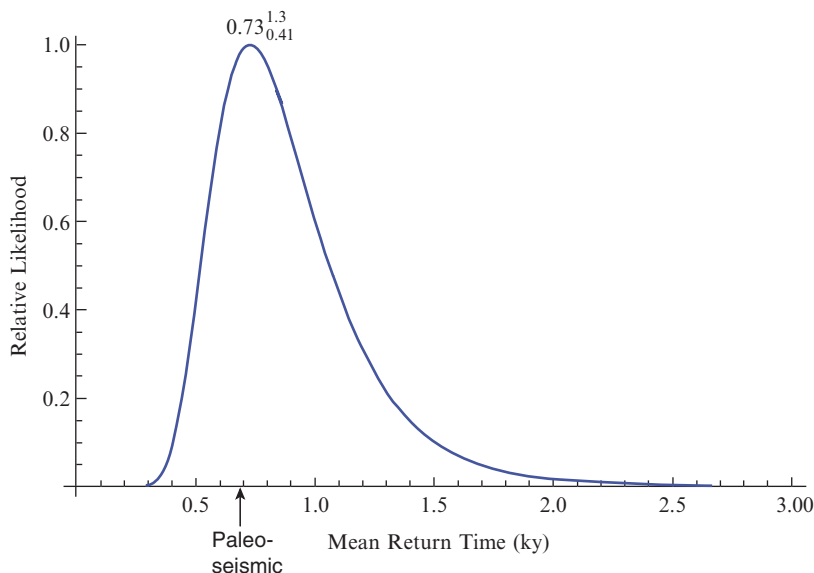


Fig. 3 Distribution of mean return times ($1/\lambda$) for debris flows in Port Valdez ($N = 5$). Blue curve: time base is 4 ka. Most likely mean return time (ka) and 95% confidence range (sub- and super-scripts) in the mean return time indicated above curve. Mean return time for paleoseismic events (0.69 ka) indicated by arrow (Carver and Plafker 2008)

earthquake (approximately 913–808 yr BP from a nearby paleoseismic site). Ryan et al. (this volume) suggest that all of the debris flows imaged in the seismic reflection data were triggered by large megathrust earthquakes younger than 4 ka. Therefore, in the statistical analysis the time period of observation is $T = 4,000$ years. The resulting distribution of mean return times is shown in Fig. 3.

The most likely mean return time of 730 years is very close to the 690 year mean return time (excluding open intervals) of megathrust earthquakes from paleoseismic data (Carver and Plafker 2008), supporting the hypothesis of Ryan et al. (this volume).

4 Sub-Events Within Landslide Complexes

Within landslide complexes such as Goleta, there are often numerous sub-events identified from geophysical data. Sub-events are defined here as individual landslides in a complex that can be distinguished using geophysical and geologic data. Two questions then arise: (1) with regard to the hazards landslides pose, do sub-events occur essentially simultaneously (e.g., with respect to the phase speed of wave propagation for tsunami hazards)? (2) Can sub-events be considered as occurring simultaneously with regard to the above empirical probability analysis (i.e., as independent events), given

the uncertainty in the age dates? The Holocene Storegga landslide complex presents an ideal case for analysis of these questions, because of the detailed observations of the complex from many authors and the large dimensions of the failure.

4.1 Case Study 3: The Holocene Storegga Landslide Complex

The Holocene Storegga landslide complex in the Norwegian Sea has been mapped in detail using multibeam bathymetry, side-scan sonar, and ROV-based bathymetry (see summary by Haflidason et al. 2005). In addition, overlapping relationships between sub-events are established from high resolution seismic reflection data, and absolute age dates of many of the sub-events have been obtained from cores.

Most sub-events are dated during the time of the main failure and have a mean of approximately 7250 yr BP (Haflidason et al. 2005). There are also as many as 10 other small sub-events that have younger age dates and are not simultaneous with the main failure. It is likely that most, if not all, of these later small failures are related to (i.e., conditional upon) the main event, owing to their location along the northern slide escarpment.

If the entire Holocene Storegga complex failed more or less synchronously, then the dates associated with the main failure can be assumed to follow a normal distribution owing to uncertainties in age-dating techniques (although the actual distribution may be more complex) (Bronk Ramsey 1998; Steier and Rom 2000). Alternatively, we might expect intervals between different sub-events to follow an exponential distribution as in the earlier examples discussed in this paper. To decide between the alternatives, we used the Shapiro-Wilk test (Shapiro and Wilk 1965; Stephens 1974) which is useful for testing empirical distributions based on small sample sizes. The test on the ensemble of failure dates associated with the main failure ($N = 45$) indicates that the normal distribution hypothesis can be rejected at a 90% confidence level. If, however, we exclude just the single youngest date from that ensemble (6,460 yr BP associated with a small sub-event near the head of the Holocene Storegga complex), then the normal distribution hypothesis cannot be rejected. Because the younger sub-events are much smaller than other sub-events that make up the main episode, it is reasonable to treat this landslide complex as a single event in geologic time. However, because of the uncertainty related to age dating, we cannot determine whether sub-events occurred alone or together with other sub-events in terms of the hazards they pose. Overlapping relationships and geotechnical arguments can aid in this determination (e.g., Locat et al., 2009).

6 Conclusions

Determining the hazards posed by submarine landslides is partly dependent on an understanding of their probability of occurrence. Assuming that multiple occurrences of large landslide complexes imaged by seismic reflection data represent

independent events, the primary probability distribution parameter that is estimated is the rate parameter (λ). Ideally, in locations where we have a sequence of landslide events, we would also have age dates of sediment that bracket each of the landslides to which we could directly apply statistical techniques developed to analyze paleoseismic sequences (Savage 1994; Ogata 1999; Parsons 2008). Presently, however, we typically only have age dates of the youngest landslide and of a seismic horizon that lies below all of the landslide events. Applying the Bayes Poisson-Gamma model to this situation, we determine the most likely ages and 95% confidence bounds for two examples: from the Santa Barbara Channel and Port Valdez. Analysis of the age dates representing the main failure episode of the Holocene Storegga landslide complex suggests that landslide complexes can be treated independently for this analysis, given uncertainty in age dating techniques. This analysis treats the Poisson probability as stationary (i.e., constant λ). However, the rate of occurrence may be linked to glacial cycle and climate change (e.g., Lee 2009; Ryan et al. this volume), that would imply a non-stationary Poisson process. More age dates of both global and local landslide sequences are needed to determine the long-term variation in λ .

Acknowledgments The authors would like to thank Hafliði Hafliðason and Uri ten Brink for their reviews of the manuscript and Holly Ryan and Mike Fisher for sharing the results of their geophysical analysis and for their comments.

References

- Borrero JC, Dolan JF and Synolakis CE (2001) Tsunamis within the eastern Santa Barbara Channel. *Geophys Res Lett* 28: 643–647.
- Brink Ramsey C (1998) Probability and dating. *Radiocarbon* 40: 461–474.
- Campbell KW (1982) Bayesian analysis of extreme earthquake occurrences. Part I. Probabilistic hazard model. *Bull Seismol Soc Am* 72: 1689–1705.
- Carver G and Plafker G (2008) Paleoseismicity and neotectonics of the Aleutian subduction zone – an overview. In: Freymueller JT, Haeussler PJ, Wesson RL and Ekström G (eds) *Active Tectonics and Seismic Potential of Alaska*. *Geophys Mon Ser* 179. Am Geophys Union, Washington, D.C., pp. 350.
- Chaytor JD ten Brink US Solow AR and Andrews BD (2009), Size distribution of submarine landslides along the U.S. Atlantic Margin, *Mar Geol* 16–27.
- Fisher MA, Normark WR, Greene HG, Lee HJ and Sliter RW (2005) Geology and tsunamigenic potential of submarine landslides in Santa Barbara Channel, southern California. *Mar Geol* 224: 1–22.
- Geist, E.L., and T. Parsons (2009), Assessment of source probabilities for potential tsunamis affecting the U.S. Atlantic Coast, *Mar Geol* 264, 98–108.
- Geist EL, Parsons T, ten Brink US and Lee HJ (2009) Tsunami Probability. In: Bernard EN and Robinson AR (eds), *The Sea* 15: 93–135. Harvard Univ Press, Cambridge, Mass.
- Hafliðason H, Lien R, Sejrup HP, Forsberg CF and Bryn P (2005) The dating and morphometry of the Storegga Slide. *Mar and Petrol Geol* 22: 123–136.
- Hutton EWH and Syvitski JPM (2004) Advances in the numerical modeling of sediment failure during the development of a continental margin. *Mar Geol* 203: 367–380.
- Lee, H. J. (2009), Timing of occurrence of large submarine landslides on the Atlantic ocean margin, *Mar Geol* 53–64.
- Locat, J., H. Lee, U. ten Brink, D. Twichell, E.L. Geist, and M. Sansoucy (2009), Geomorphology, stability and mobility of the Currituck slide, *Mar Geol* 28–40.

- Mortgat CP and Shah HC (1979) A Bayesian model for seismic hazard mapping. *Bull Seismol Soc Am* 69: 1237–1251.
- Normark WR, McGann M and Sliter RW (2004) Age of Palos Verdes submarine debris avalanche, southern California. *Mar Geol* 203: 247–259.
- Ogata Y (1999) Estimating the hazard of rupture using uncertain occurrence times of paleoearthquakes. *J Geophys Res* 104: 17 995–18 014.
- Parsons T (2008) Monte Carlo method for determining earthquake recurrence parameters from short paleoseismic catalogs: Example calculations for California. *J Geophys Res* 113: doi:10.1029/2007J–B004998.
- Parsons T and Geist EL (2009) Tsunami probability in the Caribbean region. *Pure and Applied Geophysics* 165: 2089–2116.
- Ryan HF, Lee HJ, Haeussler PJ, Alexander CR and Kayen RE (this volume) Historic and paleo-submarine landslide deposits imaged beneath Port Valdez, Alaska: implications for tsunami generation in a glacial fiord. In: Mosher DC, Shipp C, Moscardelli L, Baxter C, Chaytor J and Lee H (eds) *Submarine Mass Movements and Their Consequences IV, Advances in Natural and Technological Hazards Research*. Springer, New York.
- Savage JC (1994) Empirical earthquake probabilities from observed recurrence intervals. *Bull Seismol Soc Am* 84: 219–221.
- Shapiro SS and Wilk MB (1965) An analysis of variance test for normality (complete samples). *Biometrika* 52: 591–611.
- Steier P and Rom W (2000) The use of Bayesian statistics for ^{14}C dates of chronologically ordered samples: A critical analysis. *Radiocarbon* 42: 183–198.
- Stephens MA 1974 EDF statistics for goodness of fit and some comparisons. *J Am Stat Assoc* 69: 730–737.
- Stuiver M and Braziunas TF (1992) Modeling atmospheric ^{14}C influences and ^{14}C ages of marine samples to 10,000 BP. *Radiocarbon* 35: 137–189.
- ten Brink US, Geist EL and Andrews BD (2006) Size distribution of submarine landslides and its implication to tsunami hazard in Puerto Rico. *Geophys Res Lett* 33: doi:10.1029/2006GL026125.
- ten Brink, US, Lee HJ, Geist EL, and Twichell D (2009), Assessment of tsunami hazard to the U.S. Atlantic Coast using relationships between submarine landslides and earthquakes, *Mar Geol* 65–73.